Systematic Review

Academic Performance and Perceptions of Undergraduate Medical Students in Case-Based Learning Compared to Other Teaching Strategies: A Systematic Review with Meta-Analysis

Diogo Maia¹, Renato Andrade²,³,⁴, José Afonso⁵, Patrício Costa⁶,⁷,⁸, Cristina Valente²,³ and João Espregueira-Mendes¹,²,³,⁷,⁹,∗

¹ School of Medicine, University of Minho, 4710-057 Braga, Portugal
² Clínica Espregueira—FIFA Medical Centre of Excellence, 4350-415 Porto, Portugal
³ Dom Henrique Research Centre, 4350-415 Porto, Portugal
⁴ Porto Biomechanics Laboratory (LABIOMEP), Faculty of Sports, University of Porto, 4200-450 Porto, Portugal
⁵ Centre for Research, Education, Innovation and Intervention in Sport (CIFI2D), Faculty of Sport, University of Porto, 4200-450 Porto, Portugal
⁶ Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, 4710-057 Braga, Portugal
⁷ ICVS/3B’s—PT Government Associate Laboratory, 4710-057 Braga, Portugal
⁸ Faculty of Psychology and Education Sciences, University of Porto, 4200-135 Porto, Portugal
⁹ 3B’s Research Group—Biomaterials, Biodegradables and Biomimetics, Headquarters of the European Institute of Excellence on Tissue Engineering and Regenerative Medicine, University of Minho, 4805-017 Barco, Portugal

∗ Correspondence: espregueira@dhresearchcentre.com

Abstract: Case-based learning (CBL) is a teaching method centered on active student learning that can overcome the limitations of traditional teaching methods used in undergraduate medical education. The aim of this systematic review was to compare the effectiveness of CBL against other teaching methodologies in terms of academic performance and perceptions (intra-individual, inter-peer and student–faculty) of undergraduate medical students. Literature searches were performed using PubMed, EMBASE and Web of Science databases up to 28 April 2021. We included studies that quantitatively compared the academic performance and perception outcomes of CBL against other teaching methodologies in undergraduate medical students. The risk of bias was judged using the RoBANS tool and certainty of evidence using the GRADE framework. Meta-analyses were conducted using a random-effects model and reported as standardized mean differences (SMD) with 95% confidence intervals (CI). Of a total of 4470 records, 41 studies comprising 7667 undergraduate medical students fulfilled the eligibility criteria and were included in our systematic review. The CBL group was superior to the other teaching method groups in terms of academic performance measured by exam scores (SMD = 2.37, 95% CI 1.25–3.49, large effect, very low certainty) and interest and motivation (SMD = 0.79, 95% CI 0.13–1.44, moderate effect, very low certainty). Other academic performance or perception outcomes were not statistically different between CBL and other teaching methods when considering the pooled effect. Still, they were often superior in the CBL group for specific subgroups. CBL showed superior academic performance (especially compared to didactic lectures and tutorial-based teaching) and interest and motivation compared to other teaching methods used with undergraduate medical students. However, the certainty of evidence was very low and further studies are warranted before a stronger and more definitive conclusion can be drawn.

Keywords: case-based learning; academic performance; medical education; medicine; teaching methodologies
1. Introduction

An undergraduate medical degree presents a demanding academic challenge and should be taught using an effective method that facilitates learning. The medical curriculum should be designed and developed with the aim of meeting the learning needs of a medical student [1–6] based on predefined core learning objectives, outcomes and competencies [7–9]. Preclinical medical education has shifted over the past few decades, with teaching strategies going beyond the syllabic content and focusing also on the learning experience and the skills needed to manage clinical cases (e.g., rational and clinical thinking, decision making and problem solving) [10,11]. As medical teaching moves towards a competency-based curriculum with a focus on student performance, teaching methodologies that are student-centered and based on objectives and learning outcomes are becoming more important and increasingly popular.

There are many teaching methods that have been implemented worldwide within medical education, but the traditional method using didactic lectures is probably still the most commonly implemented [12,13]. Didactic lectures consist of the active transmission of theoretical knowledge by the teacher, while the student plays a (mostly) passive role based on listening, annotation and acceptance of the content presented by the teacher. Didactic lectures are more commonly implemented in first-year and second-year at the medical undergraduate course to provide a baseline foundation of information for subsequent years. However, this method has received much criticism in the literature [14–16], especially due to its passive nature [17] (knowledge is transmitted by and from the teacher and passively received by the students) as a teacher-centered strategy. While didactic lectures allow to communicate a large quantity of information [16,18–20], they rely on the students’ attention span [21] and their ability to memorize this large amount of knowledge [22], which is limited due to the poor performance of short-term and long-term memory for lectures over 20 min [23–25], the overload of information communicated and the lack of dynamic interaction between the student and the teacher [11,13,18,26]. More importantly, despite the benefits of lecture-style teaching, didactic lectures in undergraduate medical teaching have been questioned due to the lack of important clinical critical thinking and reasoning skills, a key factor in a medical education curriculum and in any medical career. Aiming to overcome the limitations of didactic lectures, more interactive teaching methods have been investigated and implemented, including case-based learning (CBL), team-based learning (TBL), problem-based learning (PBL), human patient simulation (HPS) and tutorial-based method (Box 1). The introduction of these student-oriented teaching methodologies into the preclinical medical curriculum grants a chief benefit of promoting clinical decision-making and problem-solving skills at an early stage of medical training [27].

CBL is a teaching method that can be defined as a form of inquiry-based learning that integrates structured and guided learning [11]. It is both a student-centered (active learning) and patient-centered (clinical cases) method [28–31]. Although several pedagogical models [32,33] can interplay in CBL, it mainly follows a constructivist model [11,32,34] (similarly to other student-centered methods) that focuses on the active role of students in creating their own knowledge by engaging with clinical cases. CBL relies on discussing clinical cases (real or hypothetical) that aim to simulate a real medical scenario and has particular relevance for medical teaching. Within the scope of medical education, the goal of CBL is to prepare medical students for their future clinical practice by challenging them with realistic clinical cases wherein they are expected to apply their theoretical background knowledge (previously acquired during preparatory classes) to clinical practice (authentic clinical cases) [11]. It promotes a more active learning style and allows the students to develop critical thinking, analytical and reasoning skills to solve clinical cases [35–37]. The teacher (facilitator) provides detailed information about the clinical case (vignettes) and stimulates a gradual discussion of the case among small groups of students [38–42]. Based on the clinical vignettes, the student has to carry out an active search for evidence-based knowledge on the topic and share their ideas with their peers on how they propose to make clinical decisions, which solutions they suggest to manage the case and to identify
key characteristics of the condition being discussed [35,43–45]. Guidelines and examples on how to prepare, structure and implement group-based CBL (including lesson plans and case studies) in undergraduate medical teaching can be found elsewhere [20,27,46–53]. At the end of the class, the clinical case is usually summarized and discussed with the teacher to consolidate key messages and the most relevant findings [54,55]. CBL is thus a versatile method in which the teacher has a less active role and should focus on guiding the students to find solutions to their queries [29,43,47,56–59] and introducing the discussion of key learning outcomes [18,22].

CBL engages students in linking the classroom-based theory of basic science and principles of clinical management [29,30,40,60,61] into their clinical practice [30,62,63]. Its role is to guide learning through clinical cases, which is of particular relevance for medical education as the medical profession is highly involved in managing patient cases. Although CBL stimulates all the learning domains under the revised Bloom’s taxonomy [64], it has a major focus on the “evaluate” and “create” domains, which are less focused in the more traditional teaching method (didactic lectures). These domains play a major role in the learning process of medical students as they prepare them for the clinical demands of their careers as well as stimulating skills that are important for their competency and performance. CBL provides students with opportunities to formulate diagnoses and delineate possible management solutions [65], as well as to understand how underlying mechanisms can relate to the diagnosis and treatment of illness [66]. Students engage in these case-based and problem-solving activities with support and feedback from their peers and experts (teachers) [67]. The preparation and discussion of clinical cases helps students to create more effective clinical thinking habits [68]. Similarly to other student-centered teaching methods (TBL, PBL and HPS, among others), CBL encourages students to autonomously develop skills in problem-solving and decision-making as well as critical and analytical thinking [69–71]. These skills are of utmost importance to medical students and will play a major role when translating knowledge into clinical practice. Although not all CBL activities must be carried out in groups, it is most commonly implemented in small groups, thereby promoting active and interpeer collaboration [63,72], which are valuable components of medical practice. As compared to PBL, the CBL method is more structured and more closely guided using predefined learning outcomes that are obtained through an inquiry-based approach [73].

Although the CBL teaching method has been implemented in the undergraduate medical degree programs of several universities, and notwithstanding the theoretical advantages already identified, the comparison of its effectiveness compared to other teaching methods has not been quantitatively systematized in the literature. A previous systematic review [11] critically analyzed CBL interventions in medical education, but searches were only performed up to 2010 and new studies have been published since then. Moreover, this previous systematic review did not provide a focused summary nor a quantitative systematization of available findings, which is needed to investigate whether the theoretical advantages are consistently seen and reproducible throughout scientific studies. Therefore, this systematic review and meta-analysis aimed to systematize the available information on CBL and fill the current knowledge gap in the literature. Our specific objective was to compare CBL against other teaching methods in terms of academic performance and perceptions (intra-individual, interpeer and student–faculty) of undergraduate medical students. The main research question was whether CBL would improve academic performance in undergraduate medical students as compared to other teaching methods, especially those that are teacher-centered (e.g., didactic lectures).
2. Materials and Methods

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [74].
2.1. Eligibility Criteria

The eligibility criteria were set according to the Participants, Interventions, Comparators, Outcomes and Study Design (PICOS) framework. Due to the lack of access to translation services, we only included studies written in English, Portuguese and Spanish.

2.1.1. Participants

We considered studies that included medical degree undergraduate students. Studies that included students from other health-allied science courses (dental medicine, nursing, veterinary medicine, psychology, physiotherapy) or those outside the health-allied field were excluded. Studies that incorporated participants in postgraduate training—such as physicians or residents—were also excluded.

2.1.2. Interventions

We included all studies that implemented the CBL teaching method as an intervention group. There are many broad definitions of the term CBL and there is no clear and widely accepted definition of CBL. We defined CBL as a teaching method centered on presenting and discussing realistic clinical cases with a group of students who play an active role in their own learning. However, it was not always clear what definition the studies used for CBL and we looked for varying definitions that studies used that could fit within the abovementioned definition. Studies that combined CBL with other complementary methods—such as PBL, TBL, Flipped Classroom (FC), digital clinical support platforms or WhatsApp group messaging—were also included. Studies that only used the PBL method as an intervention group without a main component of CBL were excluded from this systematic review, as were studies in which the CBL method was applied outside the university context.

2.1.3. Comparators

All studies that compared CBL with other teaching methods in medical education were considered for inclusion: didactic lectures, HPS, PBL, TBL, tutorial-based method, independent readings, lectures with tutorial sessions and unspecified teaching methods. While we allowed unspecified teaching methods as a comparison group (for overall analyses), we excluded them during sensitivity analyses.

2.1.4. Outcomes

The outcomes considered for inclusion in this systematic review were (i) academic performance determined through exam evaluation and (ii) perceptions about CBL (intra-individual, interpeer, student–faculty and other relevant perceptions. All outcomes had to be provided with quantitative data analysis to be included in the analysis.

2.1.5. Study Design

We included studies that employed study designs that allowed the comparison of CBL with other teaching methodologies: randomized or quasi-randomized studies, non-randomized cohort studies, case–control studies, crossover and cross-sectional studies. While exclusively qualitative studies (e.g., only based on structured interviews) were excluded, we allowed qualitative and mixed-methods studies that quantified results.

2.2. Search Strategy

We undertook a comprehensive search on PubMed, EMBASE and Web of Science databases to identify studies that compared CBL against other teaching methodologies in terms of academic performance and/or perceptions of CBL among undergraduate medical students. The search strategy was applied from database inception up until 28 April 2021 (Table S1). The reference lists of reviews that were relevant to this topic and studies that satisfied the eligibility criteria were manually searched for other potentially eligible studies not identified via the database searches.
2.3. Study Selection

All records resulting from database searches were exported to the “Mendeley Desktop” (Elsevier Solutions), where duplicate studies were automatically removed and confirmed through a manual search. Two authors (D.M., R.A.) independently examined all titles and abstracts found to identify potentially relevant studies for inclusion. Relevant studies were read in full to determine whether they met the eligibility criteria. Disagreements were resolved by consensus.

2.4. Data Collection and Extraction

All data were extracted from the included studies by one author (D.M.) and rechecked by another two authors (R.A., J.A.). Data were collected in an Excel file and organized into different subsections: characteristics of the study (citation, year of study, country in which the study was carried out, study design, curricular area and curricular year), characteristics of the population (total sample and divided by CBL and comparison(s) group(s), and sample number for each curricular year) and the outcome measures comparing CBL with other teaching methods. Teaching methodologies were characterized according to models, strategies and methods [75]. Outcome measures were grouped according to clusters of measures: “academic performance”, “intra-individual student perceptions”, “interpeer perceptions”, “student–faculty perceptions” and “other perceptions”.

2.5. Data Management

Means and standard deviations were extracted for each outcome measure. When it was not possible to extract the standard deviation, the methods suggested by the Cochrane Handbook [34] were employed to impute the standard deviation using the standard error or the confidence interval associated with the statistical significance value (p value) between the two groups (experimental and control). In cases where the mean and standard deviation values were represented in figures, WebPlotDigitizer software [35] was used to extract these values.

For the quantitative synthesis, when a study presented more than one outcome measure within the same cluster of measures, the metric that best represented the outcome being evaluated was prioritized. For example, if a study reported final and other intermediate academic performance exam results, the final academic performance exam was prioritized. When the same study presented several evaluation timepoints, including groups within the same population of students, the sample size of each group was divided by the number of evaluations for each timepoint/evaluation, thus avoiding overlapping and over-representation of the same population in the analysis. Similarly, for crossover studies where the sample size was overlapped for both groups, the sample size for each group was divided in half. In studies that had more than one comparison group, the sample size of the CBL or control group (the one being compared twice) was proportionally divided by the sample size of the other comparison groups.

2.6. Risk of Bias

Two authors (R.A., J.A.) judged the risk of bias using the Risk of Bias Assessment tool for Non-randomized Studies (RoBANS) tool [76]. The RoBANS is a validated tool used to assess the risk of bias in nonrandomized studies, comprising six domains of bias, including the selection of participants, confounding variables, measurement of exposure, blinding of outcome assessment, incomplete outcome data and selective outcome reporting. Each domain is judged as low risk, high risk or unclear risk of bias. There is no specific tool to appraise the risk of bias in medical education studies. We decided to use the RoBANS tool, which evaluates several key domains of risk of bias, and adapted the judgment criteria to the topic being studied (Table S2).
2.7. Data Synthesis

We employed quantitative synthesis methods to allow a head-to-head comparison between CBL and the other teaching methodologies. For the quantitative syntheses, we conducted a meta-analysis using Rstudio 3.3.1 software (Rstudio, Boston, MA, USA) and the “devtools”, “meta” and “metafor” packages. Meta-analysis was performed when at least three studies/entries reported the same outcome measure (i.e., within the same cluster) comparing a CBL group with any other teaching methodology. Outcomes that were not reported in a minimum of three studies and/or not homogenously evaluated across studies were excluded from the meta-analysis. The meta-analysis framework provided a head-to-head comparison between CBL and other teaching methods (didactic lectures, TBL, PBL, HPS and tutorial method) for each available outcome. The main analyses were stratified according to the different outcome measures for all curriculum areas. Then, subgroup analyses were conducted according to the curriculum area to investigate the effect size for each outcome within each area.

The random-effects model (Sidik–Jonkman estimator with Hartung–Knapp adjustment [77]) was used to attenuate the variations in the methodology and evaluation methods across studies and within the same outcome measures (methodological heterogeneity). This model is appropriate when there is considerable methodological heterogeneity and/or between-study heterogeneity in the true effects. The impact of statistical heterogeneity was established using the $I^2$ and interpreted as unimportant (<50%), moderate (50–75%) or high (>75%) [78]. Quantitative syntheses were expressed as the standardized mean differences (SMDs) and their 95% confidence intervals (CI). The magnitude of the SMDs were interpreted as high ($\geq 0.8$), moderate (0.5 to 0.79) or weak (0.2 to 0.49) [79].

Sensitivity analyses were performed to explore potential sources of heterogeneity and analyze their influence on the pooled statistical heterogeneity and pooled effect size. Sensitivity analyses involved rerunning the main analyses, but removing studies with characteristics that could influence the results: (i) studies that had an outlier effect (the 95% CI of the study did not cross the 95% CI of the pooled effect); (ii) studies with a crossover design; (iii) studies in which the experimental group differed from traditional CBL (e.g., with an added component to CBL); (iv) studies in which academic performance was assessed for a focused topic (e.g., diabetes) or did not reflect overall (final) academic performance. We also conducted two additional sensitivity analyses comparing the results of CBL against didactic lectures and against the tutorial method (the only comparison groups with enough studies for single-method analyses). A minimum of two studies had to remain after exclusions in order to perform the sensitivity analysis.

The risk of publication bias was statistically analyzed using the Egger’s test [80] and by inspecting the funnel plots when the meta-analysis included 10 or more studies. However, it should be noted that publication bias should not be automatically inferred, as many other causes can explain funnel plot asymmetry [81–83]. Trim-and-fill analysis [84] was also performed to verify whether there was an adjustment in the effect size (SMD) when correcting for the risk of publication bias.

2.8. Certainty of Evidence

Two authors (R.A., C.V.) judged the certainty of evidence, which was summarized using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach [85]. Certainty of evidence was graded as high, moderate, low or very low certainty. Certainty was downgraded if there were concerns related to risk of bias, inconsistency, imprecision, indirectness or risk of publication bias. When the meta-analysis included fewer than 10 studies, risk of publication bias was not considered for downgrading the certainty of evidence.
3. Results

3.1. Search Results and Study Characteristics

Database and manual searches yielded 4470 records. After screening the titles and abstracts, 239 studies were selected for full-text analysis, of which 41 studies met the eligibility criteria and were included in the systematic review [13,18,20,27,37,47–53,55,69,86–112]. The reasons for excluding the remaining 198 studies are reported in the PRISMA flow diagram (Figure 1).

Figure 1. PRISMA flow diagram for searches and study selection.

The majority of the included studies were published within the last 10 years (36 out of 41 studies). Geographical distribution showed that studies were mostly conducted on the Asian continent (27 studies), followed by the North American continent (10 studies), and with only a few conducted on the European continent (4 studies).

There were 11 studies with a crossover design, 10 studies with a cross-sectional design (post), 16 studies with a pre–post design and 4 with an intermediate–post design (Table 1).

Table 1. Characteristics of included studies, study populations and outcome measures.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Country</th>
<th>Curriculum Area</th>
<th>Sample</th>
<th>Curriculum Year</th>
<th>Outcome Measures</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adiga et al., 2011 [113]</td>
<td>India</td>
<td>Biochemistry</td>
<td>CBL: 110 DL: 110</td>
<td>2nd year: 110</td>
<td>Academic performance</td>
<td>Post (crossover)</td>
</tr>
<tr>
<td>Ahmad et al., 2017 [37]</td>
<td>Kuwait</td>
<td>Medical Education</td>
<td>CBL: 113 DL: 146</td>
<td>5th and 6th years: 259</td>
<td>Intra-individual perceptions</td>
<td>Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Student–faculty perceptions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Other perceptions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Other perceptions</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Country</th>
<th>Curriculum Area</th>
<th>Sample</th>
<th>Curriculum Year</th>
<th>Outcome Measures</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bhardwaj et al., 2015 [87]</td>
<td>India</td>
<td>Community Medicine Microbiology Internal Medicine Obstetrics and Gynecology</td>
<td>CBL: 90 DL: NR</td>
<td>3rd year: 90 4th year: NR</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Carrasco et al., 2018 [88]</td>
<td>USA</td>
<td>Microbiology Immunology Oncology</td>
<td>CBL: 23 TBL: 23</td>
<td>1st year: 23</td>
<td>Other perceptions</td>
<td>Inter–Post (crossover)</td>
</tr>
<tr>
<td>Ciraj et al., 2010 [13]</td>
<td>India</td>
<td>Microbiology Immunology Oncology</td>
<td>CBL: 166 DL: 166</td>
<td>2nd year: 166</td>
<td>Academic performance</td>
<td>Inter–Post</td>
</tr>
<tr>
<td>Diwan et al., 2017 [90]</td>
<td>India</td>
<td>Endocrinology</td>
<td>CBL: 13 Tutorial method: 13</td>
<td>1st year: 26</td>
<td>Academic performance</td>
<td>Pre–Post (crossover)</td>
</tr>
<tr>
<td>Fortun et al., 2017 [91]</td>
<td>USA</td>
<td>Molecular Biology Biochemistry Medical Genetics</td>
<td>CBL: 154 Independent readings: 97</td>
<td>1st year: 130</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Grover et al., 2020 [92]</td>
<td>India</td>
<td>Pathology</td>
<td>CBL + WhatsApp: 35 DL: 35</td>
<td>2nd year: 70</td>
<td>Academic performance</td>
<td>Pre–Post (crossover)</td>
</tr>
<tr>
<td>Hansen et al., 2005 [93]</td>
<td>USA</td>
<td>Obstetrics and Gynecology</td>
<td>CBL: 90 DL: 32</td>
<td>3rd year: 122</td>
<td>Intra-individual perceptions Student–faculty perceptions</td>
<td>Post (crossover)</td>
</tr>
<tr>
<td>Hashim et al., 2015 [94]</td>
<td>Pakistan</td>
<td>Pathology</td>
<td>CBL: 144 Tutorial method: 144</td>
<td>4th year: 144</td>
<td>Intra-individual perceptions Interpeer perceptions Student–faculty perceptions</td>
<td>Pre–Post (crossover)</td>
</tr>
<tr>
<td>Joshi et al., 2014 [95]</td>
<td>India</td>
<td>Biochemistry</td>
<td>CBL: 30 DL: 30</td>
<td>1st year: 60</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Kamat et al., 2012 [96]</td>
<td>India</td>
<td>Pharmacology</td>
<td>CBL: 96 DL: 83</td>
<td>2nd year: 179</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Kaur et al., 2020 [97]</td>
<td>India</td>
<td>Pharmacology</td>
<td>CBL: 47 DL: 47</td>
<td>2nd year: 94</td>
<td>Academic performance Other perceptions</td>
<td>Post (crossover)</td>
</tr>
<tr>
<td>Lee et al., 2013 [98]</td>
<td>Taiwan</td>
<td>Nuclear Medicine</td>
<td>CBL: 36 DL: 34</td>
<td>5th year: 70</td>
<td>Intra-individual perceptions Interpeer perceptions Student–faculty perceptions</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Ma et al., 2016 [99]</td>
<td>China</td>
<td>Immunology</td>
<td>CBL: 85 DL: 85</td>
<td>NR: 85</td>
<td>Academic performance</td>
<td>Inter–Post (crossover)</td>
</tr>
<tr>
<td>Citation</td>
<td>Country</td>
<td>Curriculum Area</td>
<td>Sample</td>
<td>Curriculum Year</td>
<td>Outcome Measures</td>
<td>Design</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------------</td>
<td>------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Maas et al., 2018 [100]</td>
<td>USA</td>
<td>Oncology</td>
<td>TB-CBL: 352 DL: 231</td>
<td>6th year: 583</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Montaldo et al., 2013 [101]</td>
<td>Chile</td>
<td>Semiology Internal Medicine</td>
<td>CBL: 37 DL + tutorial with patients: 27</td>
<td>3rd year: 64</td>
<td>Academic performance Intra-individual perceptions</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Nair et al., 2013 [50]</td>
<td>India</td>
<td>Biochemistry</td>
<td>CBL: 50 DL: 50</td>
<td>1st year: 100</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Palappallil et al., 2019 [103]</td>
<td>India</td>
<td>Pharmacology</td>
<td>CBL: 44 Control 1 (TBL): 50 Control 2 (DL): 49</td>
<td>2nd year: 145</td>
<td>Academic performance Intra-individual perceptions Student–faculty perceptions Other perceptions</td>
<td>Post</td>
</tr>
<tr>
<td>Panja et al., 2013 [104]</td>
<td>India</td>
<td>Physiology</td>
<td>CBL: 150 Tutorial method: 150</td>
<td>1st year: 300</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Patil et al., 2016 [105]</td>
<td>India</td>
<td>Microbiology</td>
<td>CBL: 29 Tutorial method: 28</td>
<td>2nd year: 58</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Rajan et al., 2016</td>
<td>India</td>
<td>Neurology</td>
<td>CBL: 31 DL: 29</td>
<td>1st year: 60</td>
<td>Academic performance</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Schwartz et al., 2007 [51]</td>
<td>USA</td>
<td>Emergency Medicine</td>
<td>CBL: 52 HPS: 50</td>
<td>4th year: 102</td>
<td>Academic performance</td>
<td>Post</td>
</tr>
<tr>
<td>Surapaneni et al., 2010 [108]</td>
<td>India</td>
<td>Biochemistry</td>
<td>CBL: 75 DL: 75</td>
<td>1st year: 150</td>
<td>Academic performance</td>
<td>Post</td>
</tr>
<tr>
<td>Tate et al., 2014 [109]</td>
<td>India</td>
<td>Microbiology</td>
<td>CBL: 78 DL: 78</td>
<td>2nd year: 78</td>
<td>Academic performance</td>
<td>Pre–Post (crossover)</td>
</tr>
<tr>
<td>Vedi et al., 2021 [110]</td>
<td>India</td>
<td>Anatomy Physiology</td>
<td>CBL: 43 DL: 80</td>
<td>1st year: 150</td>
<td>Academic performance</td>
<td>Inter–Post</td>
</tr>
<tr>
<td>Vora et al., 2015 [111]</td>
<td>India</td>
<td>Pharmacology</td>
<td>CBL: 34 DL: 34</td>
<td>3rd year: 68</td>
<td>Academic performance</td>
<td>Post</td>
</tr>
<tr>
<td>Wallany et al., 2019 [27]</td>
<td>USA</td>
<td>Internal Medicine</td>
<td>CBL: 29 Unspecified group: 35</td>
<td>1st year: 18 2nd year: 46</td>
<td>Intra-individual perceptions</td>
<td>Pre–Post</td>
</tr>
<tr>
<td>Yang et al., 2021 [112]</td>
<td>China</td>
<td>Nephrology</td>
<td>FC + CBL: 31 DL: 31</td>
<td>4th year: 62</td>
<td>Intra-individual perceptions Interpeer perceptions Other perceptions</td>
<td>Post</td>
</tr>
</tbody>
</table>

**Legend:** CBL—Case-based learning, DL—Didactic Lectures, FC—Flipped Classroom, HPS—Human Patient Simulation, NR—Not reported, OSCE—Objective structured clinical examination, PBL—Problem-Based Learning, TBL—Team-Based Learning, USA—United States of America. * The total number of students evaluated in the study was not clear.
3.2. Risk of Bias

All but one of the studies [51] were judged to have a high risk of bias in at least one domain (Figure 2). Around half of the studies (k = 21, 51%) were judged to have a high risk of selection bias due to patient selection, and more than half (k = 26, 63%) due to uncontrolled confounding variables. Nearly all studies (k = 37, 90%) were judged to have a high risk of performance bias due to measurement of exposure. Most studies were judged to have either high (k = 26, 63%) or unclear (k = 13, 32%) risk of detection bias. High risk of attrition bias was identified in nearly half of the studies (k = 19, 46%). No study was judged to have a low risk of selective reporting bias, with 29% (k = 12) of studies being judged to have a high risk and the remaining 71% (k = 29) an unclear risk.

![Figure 2. Risk of bias judgement: (a) study-level; (b) overall summary of all included studies. * The total number of students evaluated in the study was not clear.](image-url)
3.3. Population Characteristics

A total of 7332 undergraduate medical students were included (Table 1), of which 3168 were from the preclinical (1st to 3rd year) and 3880 from the clinical years (4th to 6th year). The curriculum year was impossible to distinguish for 199 students (between the 3rd and 4th year) and another 62 students (between the 3rd and 6th year).

3.4. Characteristics of CBL and Other Teaching Methodologies

There were 41 studies implementing CBL, 29 studies implementing didactic lectures, 3 studies implementing TBL, 4 studies implementing PBL, 1 study implementing HPS, 5 studies implementing the tutorial method and 1 study implementing independent readings. The characterization of these teaching methodologies according to their models, strategies and methods is described in Table S3.

There were 4639 students in the CBL groups and 3938 students in non-CBL comparison groups (Table 1). The comparison groups were taught using different teaching methods: 1547 students were in didactic lectures groups, 335 students in tutorial-based method groups, 745 students in unspecified teaching method groups, 97 students in independent readings groups, 73 students in TBL groups, 50 students in HPS groups, 27 students in groups receiving lectures with tutorial sessions and 64 students in PBL groups.

The most common curriculum areas investigated were biochemistry (7 studies), microbiology (5 studies), pharmacology (4 studies) and emergency medicine, general surgery, internal medicine, immunology and physiology (3 studies each).

3.5. Comparison of CBL against Other Teaching Methodologies

The outcomes were clustered into the same domains and grouped into academic performance and perceptions. Of the 45 different outcomes collected, 6 outcomes (13.3%) were related to academic performance and 39 outcomes (86.7%) to the different perceptions (Table S4). Within the academic performance domain, 30 studies reported exam scores, 3 studies reported objective structured clinical examination (OSCE) scores and the remaining four outcomes were reported in only 1 study. Perceptions were grouped into four domains, of which the intra-individual student perceptions were the most commonly reported (13 studies). Other perceptions included interpeer (6 studies), student–faculty (8 studies) and other perceptions (5 studies).

3.5.1. Academic Performance

Within the academic performance domain, only two outcomes—exam scores (k = 40) and OSCE scores (k = 3)—were eligible for meta-analysis. The forest plots are shown in Figure S5. The CBL group showed superior academic performance (exam scores) compared to other teaching methodology groups (SMD = 2.37, 95% CI 1.25 to 3.49; large effect, very low certainty). Subgroup analysis showed that the effect size was larger for the otorhinolaryngology subgroup (SMD = 11.78, 95% CI 9.65 to 13.92; large effect). Conversely, there were no significant differences between CBL and other teaching methodologies for the general surgery, microbiology, endocrinology, pharmacology, physiology and molecular biology/biochemistry/genetic medicine subgroups (Table 2). The impact of heterogeneity was high (94%) and significant for the pooled effect and the subgroup analysis, except for the otorhinolaryngology and general surgery subgroups. The funnel plot showed asymmetric effect sizes (Figure S6), and there was a potential risk of publication bias (Egger’s test, t = 5.37, p < 0.001). The trim-and-fill analysis (11 entries added) did not impact the degree of heterogeneity (Figure S6), but the effect size was no longer significant when corrected for publication bias (SMD = 0.66, 95% CI −0.77 to 2.10). Sensitivity analyses revealed that the impact of heterogeneity remained significant except when excluding studies with an outlier effect. In all sensitivity analyses, the effect size remained large and significant. When compared in isolation with didactic lectures or the tutorial method, CBL remained significantly superior (Table S7). On the other hand, academic performance measured using
OSCE scores was not significantly different between CBL and other teaching methodologies (Table 2).

**Table 2.** Quantitative synthesis of the comparison of CBL and other teaching methodologies.

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>CURRICULUM AREA</th>
<th>k, n</th>
<th>I², p Value</th>
<th>SMD (95% CI)</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exam scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>POOLED EFFECT</strong></td>
<td></td>
<td>40,4263</td>
<td>94%, p &lt; 0.001</td>
<td>2.37 (1.29 to 3.49)</td>
<td></td>
</tr>
<tr>
<td>Biochemistry</td>
<td>11,886</td>
<td>94%, p &lt; 0.001</td>
<td>3.32 (1.44 to 5.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Surgery</td>
<td>2,297</td>
<td>0%, p = 0.78</td>
<td>0.29 (–0.13 to 0.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otorhinolaryngology</td>
<td>4,96</td>
<td>0%, p = 0.70</td>
<td>11.78 (9.65 to 13.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microbiology</td>
<td>4,545</td>
<td>85%, p &lt; 0.001</td>
<td>0.79 (–0.016 to 1.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endocrinology</td>
<td>3,396</td>
<td>80%, p &lt; 0.001</td>
<td>0.20 (–2.00 to 2.41)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molecular Biology/Biochemistry/Medical Genetics</td>
<td>2,251</td>
<td>82%, p = 0.02</td>
<td>1.03 (–4.55 to 6.61)</td>
<td>⬤⬤⬤⬤ *†††</td>
<td></td>
</tr>
<tr>
<td>Pathology</td>
<td>1,70</td>
<td>NA</td>
<td>1.56 (1.02 to 2.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharmacology</td>
<td>5,367</td>
<td>90%, p &lt; 0.001</td>
<td>–0.07 (–1.10; 0.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiology</td>
<td>3,427</td>
<td>74%, p = 0.02</td>
<td>0.83 (–0.03 to 1.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunology</td>
<td>1,84</td>
<td>NA</td>
<td>0.72 (0.28 to 1.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oncology</td>
<td>1,583</td>
<td>NA</td>
<td>0.91 (0.73 to 1.08)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathology/Intern Medicine</td>
<td>1,64</td>
<td>NA</td>
<td>0.86 (0.34 to 1.38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiology</td>
<td>1,135</td>
<td>NA</td>
<td>1.58 (1.19 to 1.97)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nephrology</td>
<td>1,62</td>
<td>NA</td>
<td>4.41 (3.47 to 5.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OSCE</strong></td>
<td></td>
<td>3,1988</td>
<td>46%, p = 0.16</td>
<td>0.30 (–0.07 to 0.67)</td>
<td>⬤⬤⬤⬤ *§</td>
</tr>
<tr>
<td>Pathology</td>
<td>3,1988</td>
<td>46%, p = 0.16</td>
<td>0.30 (–0.07 to 0.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharmacology</td>
<td>4,654</td>
<td>97%, p &lt; 0.001</td>
<td>0.87 (–0.49 to 2.23)</td>
<td></td>
<td>⬤⬤⬤⬤ *††§</td>
</tr>
<tr>
<td>Physiology</td>
<td>1,70</td>
<td>NA</td>
<td>0.98 (0.49 to 1.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunology</td>
<td>1,170</td>
<td>NA</td>
<td>0.31 (–0.16 to 0.78)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oncology</td>
<td>1,374</td>
<td>NA</td>
<td>0.15 (–0.15 to 0.45)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endocrinology</td>
<td>1,374</td>
<td>NA</td>
<td>2.01 (1.75 to 2.27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Self-learning</strong></td>
<td></td>
<td>3,276</td>
<td>96%, p &lt; 0.001</td>
<td>2.76 (–2.42 to 7.94)</td>
<td>⬤⬤⬤⬤ *†§</td>
</tr>
<tr>
<td>Pathology</td>
<td>1,70</td>
<td>NA</td>
<td>1.01 (0.51 to 1.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunology</td>
<td>1,344</td>
<td>NA</td>
<td>2.26 (1.99 to 2.53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nephrology</td>
<td>1,62</td>
<td>NA</td>
<td>5.12 (4.07 to 6.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Critical thinking</strong></td>
<td></td>
<td>3,276</td>
<td>96%, p &lt; 0.001</td>
<td>2.76 (–2.42 to 7.94)</td>
<td>⬤⬤⬤⬤ *†§</td>
</tr>
<tr>
<td>Pathology</td>
<td>1,70</td>
<td>NA</td>
<td>1.01 (0.51 to 1.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunology</td>
<td>1,344</td>
<td>NA</td>
<td>2.26 (1.99 to 2.53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nephrology</td>
<td>1,62</td>
<td>NA</td>
<td>5.12 (4.07 to 6.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Satisfaction with the teaching method</strong></td>
<td></td>
<td>4,354</td>
<td>95%, p &lt; 0.001</td>
<td>1.31 (–1.60 to 4.23)</td>
<td>⬤⬤⬤⬤ *†§</td>
</tr>
<tr>
<td>Otorhinolaryngology</td>
<td>1,100</td>
<td>NA</td>
<td>0.39 (0.00 to 0.79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuclear Medicine</td>
<td>1,70</td>
<td>NA</td>
<td>0.31 (–0.16 to 0.78)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nephrology</td>
<td>1,62</td>
<td>NA</td>
<td>4.12 (3.22 to 5.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obstetrics and Gynecology</td>
<td>1,122</td>
<td>NA</td>
<td>0.55 (0.15 to 0.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time allotted for classes</strong></td>
<td></td>
<td>4,565</td>
<td>61%, p = 0.05</td>
<td>0.15 (–0.26 to 0.57)</td>
<td>⬤⬤⬤⬤ *§</td>
</tr>
<tr>
<td>Immunology</td>
<td>1,170</td>
<td>NA</td>
<td>–0.07 (–0.37 to 0.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharmacology</td>
<td>2,143</td>
<td>0%, p = 1.00</td>
<td>0.05 (0.05 to 0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Education</td>
<td>1,252</td>
<td>NA</td>
<td>0.46 (0.20 to 0.71)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Cont.

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>CURRICULUM AREA</th>
<th>k, n</th>
<th>I², p Value</th>
<th>SMD (95% CI)</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interest and motivation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POOLED EFFECT</td>
<td>6, 841</td>
<td>90%, p &lt; 0.001</td>
<td>0.79 (0.13 to 1.44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunology</td>
<td>1, 170</td>
<td>NA</td>
<td></td>
<td>0.46 (0.16 to 0.77)</td>
<td></td>
</tr>
<tr>
<td>Endocrinology</td>
<td>1, 344</td>
<td>NA</td>
<td></td>
<td>1.29 (1.06 to 1.52)</td>
<td></td>
</tr>
<tr>
<td>Nephrology</td>
<td>1, 62</td>
<td>NA</td>
<td></td>
<td>1.51 (0.94 to 2.08)</td>
<td></td>
</tr>
<tr>
<td>Obstetrics and Gynecology</td>
<td>1, 122</td>
<td>NA</td>
<td></td>
<td>−0.23 (−0.63 to 0.17)</td>
<td></td>
</tr>
<tr>
<td>Pharmacology</td>
<td>2, 143</td>
<td>0%, p = 0.99</td>
<td>0.87 (0.85 to 0.89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interpeer interaction in team work</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POOLED EFFECT</td>
<td>5, 716</td>
<td>97%, p &lt; 0.001</td>
<td>1.45 (−0.34 to 3.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathology</td>
<td>1, 70</td>
<td>NA</td>
<td></td>
<td>0.92 (0.43 to 1.42)</td>
<td></td>
</tr>
<tr>
<td>Nuclear Medicine</td>
<td>1, 70</td>
<td>NA</td>
<td></td>
<td>0.91 (0.42 to 1.40)</td>
<td></td>
</tr>
<tr>
<td>Immunology</td>
<td>1, 170</td>
<td>NA</td>
<td></td>
<td>−0.12 (−0.42 to 0.19)</td>
<td></td>
</tr>
<tr>
<td>Nephrology</td>
<td>1, 62</td>
<td>NA</td>
<td></td>
<td>3.75 (2.90 to 4.59)</td>
<td></td>
</tr>
<tr>
<td>Endocrinology</td>
<td>1, 344</td>
<td>NA</td>
<td></td>
<td>1.95 (1.70 to 2.21)</td>
<td></td>
</tr>
</tbody>
</table>

Legend: OSCE—Objective Structured Clinical Examination; k—number of intervention comparisons (note that for a single study more than one comparison can be represented); I²—heterogeneity; SMD—standardized mean difference; CI—confidence intervals. * Downgraded two levels due to very serious study limitations (high risk of bias in at least three domains); † Downgraded one level due to serious inconsistency (substantial and significant heterogeneity and/or 95% CI did not overlap); †† Downgraded two levels due to very serious inconsistency (substantial and significant heterogeneity and 95% CI did not overlap with pooled effect for at least two studies; ‡ Downgraded one level due to serious indirectness (surrogate outcomes or heterogenous methods of CBL/controls or evaluation); § Downgraded one level due to serious imprecision (wide 95% CIs crossing the 0.5 or −0.5 SMD, and/or small sample size n < 800); §§ Downgraded two levels due to very serious imprecision (very wide 95% CIs crossing both the 0.5 and −0.5 SMD, regardless of sample size). ⭐⭐⭐⭐—very low certainty of evidence.

3.5.2. Intra-Individual Student Perceptions

Five outcomes within the intra-individual student perceptions category were eligible for meta-analysis. The outcome “interest and motivation” was significantly superior in the CBL group (SMD = 0.79, 95% CI 0.13 to 1.44; moderate effect, very low certainty). The remaining four outcomes were not significantly different between CBL and other teaching methodology groups when considering the pooled effect (very low certainty). Subgroup analyses showed a significantly superior effect of CBL for most curricular areas. The impact of heterogeneity was generally high (61% to 96%) and significant, and usually remained unchanged after sensitivity analyses. The effect size did not change with sensitivity analyses, except for the “interest and motivation” outcome, which was no longer significant after removing studies with heterogenous CBL or control groups, but remained significant when removing studies with a crossover design. Forest plots are shown in Figure S5.

3.5.3. Interpeer Perceptions

Only one outcome assessed interpeer perceptions and was eligible for meta-analysis. The CBL and control groups were not statistically different for the “interpeer interaction in team work” outcome (very low certainty). However, the effect size was significantly superior in the CBL groups for all subgroups except for the immunology curriculum area (Table 2). The impact of heterogeneity and effect size remained unchanged after sensitivity analyses (Table S6). Forest plots are shown in Figure S5.

4. Discussion

4.1. Is CBL Effective for Undergraduate Medical Teaching?

Our systematic review with meta-analysis showed that there is a very low certainty of evidence that CBL is significantly superior to other teaching methodologies (especially when compared to didactic lectures or tutorial method) in improving academic performance
measured by exam scores. Although a similar significant effect should have been seen in OSCE scores, which reflect clinical cases in real-world scenarios, the meta-analysis of this outcome only accounted for two studies and may thus be unrepresentative.

CBL was also significantly superior at enhancing student interest and motivation. Interestingly, other important student-centered features related to CBL, such as critical thinking and self-learning, were not significantly different in the groups exposed to other teaching methodologies. However, subgroup analysis showed that the CBL group scored significantly higher in critical thinking for all curriculum areas and within all included studies, suggesting that these skills may be improved when using CBL. During the implementation of CBL—as a teaching method oriented towards inquiry and critical thinking—there are different epistemological and theoretical approaches (such as constructivism, cognitivism and behaviorism) [11,32–34,114,115] that can interplay to promote and enhance the skills of decision-making, problem-solving and critical and analytical thinking. A constructivist approach is foundational to CBL because it emphasizes the active role of students in creating their own knowledge (discovery learning) while engaging with clinical cases and thus building their own understanding of medical concepts and procedures. Within a constructivist approach, CBL can be seen as more effective at developing critical thinking skills than didactic lectures as it allows students to make connections between what they have learned and real-world scenarios. Social constructivism is also at play during group-based CBL where construction of knowledge occurs through social interaction and communication while working together and discussing the clinical cases. CBL can incorporate elements of cognitivism by developing the cognitive abilities of students by requesting them to process and apply clinical information for them to diagnose and suggest appropriate treatments for the assigned clinical case. For this purpose, students apply their existing cognitive models (mental representations of knowledge) to understand the clinical information presented—which may include recognizing patterns in the patient’s symptoms or drawing connections between the patient’s medical history and their current condition—and thus forming hypotheses about potential diagnoses and treatments. While working through the case, students may refine or expand their existing cognitive models as they learn new information and gain new insights into their decision-making process. By working with real-life clinical cases and applying their knowledge to new situations, students are able to build their own understanding of medical concepts, rather than simply memorizing information. CBL can also be seen through the lens of behaviorism theory (to a lesser extent than didactic lectures), as learning occurs as a result of the students’ response to clinical cases and can be observed and measured by changes in their behavior. By receiving feedback on their performance when making clinical decisions, students are able to understand the consequences of their behavior and adjust it accordingly. Repeated exposure to different realistic clinical cases and feedback can help students to reinforce and strengthen their critical thinking skills. These theoretical learning approaches can blend in CBL, combining into a dynamic process of connectivism and interbehaviorism wherein students’ understanding of clinical cases is shaped by connecting internal (previous experiences and knowledge) and external factors (active construction of meaning that takes place during CBL). By engaging with clinical cases, students are able to actively construct their own understanding of clinical cases while also being influenced by the feedback they receive from their peers and teachers. An experiential learning approach can also be considered as, during the constructivist process, students go through a process of experience, reflection and action (especially if using real patient cases) and can reflect on their own experiences and feedback received. These approaches play a role during CBL, but are also common in other student-centered methods as PBL, TBL and HPS, which promote critical thinking and other related reasoning skills. These features will play a critical role in the students’ future careers as medical doctors as they implement and interpret new knowledge to solve real clinical cases, which is often focused CBL [35,36,45,116–119].

Our findings underpin an important and relevant effect of CBL and suggest that this teaching method can be successfully implemented in undergraduate medical curricula.
These findings should, however, be interpreted with caution. The certainty of evidence was very low for all outcomes due to several limiting factors, including risk of bias, inconsistency, imprecision and indirectness. Both main and subgroup analyses displayed a high impact of heterogeneity for most outcomes, even after sensitivity analysis to exclude studies with potentially confounding factors related to study design and teaching methodologies. However, it should be noted that no meta-analysis showed a tendency or a significant effect for other teaching methodologies compared to CBL. There may have been different potential confounding factors that could not be controlled or analyzed in our sensitivity analyses and that might have influenced the development of important student-centered skills such as self-learning and critical thinking. These student-centered skills, besides being inherently influenced by intra-individual student characteristics (either behavioral, motivational or psychological) [115,120], can be confounded by external factors. Although we excluded studies where the CBL method was applied outside the university context, students might have developed or enhanced these skills by attending extracurricular courses outside the intended curricula or even through previous experience with other teaching methods (e.g., PBL) that might have provided some advantages in developing critical thinking skills. However, such instances were not reported in most of the included studies and could not be controlled or accounted for in our analyses. The socioeconomic and educational factors of the family of origin are also known to influence students’ personal and professional development and performance [121–124], which may have confounded our results; however, these factors could also not be controlled or accounted for in our analyses. Lastly, the characteristics, profile and charisma of teachers and the strategies they used to implement CBL (or any of the other teaching methods) may have exerted an important influence on the experience of learning and ultimately affected the students’ academic performance and their experience and perceptions of CBL and other teaching methods.

CBL (and other student-centered teaching methods) have not come to completely supersede the traditional method of teaching. Didactic lectures still have a place in medical education as they convey a large amount of basic knowledge to large student classes in a simple and time-efficient fashion. This teaching method may be more suitable for the basic years of medical curriculum where the teacher needs to transmit a huge amount of basic and fundamental information that students must acquire and which will be foundational in later years. This early exposure to fundamental knowledge will allow to enhance the assimilation of knowledge through the discussion of cases [125]. Moreover, teaching methods do not need to be implemented in a standalone fashion, but can rather be combined to allow the advantages of each method to complement other methods. Didactic lectures can thus be combined and complemented with one or more active teaching methods (CBL, TBL, PBL, etc.) to promote students’ autonomy and responsibility in learning while promoting valuable real-world cognitive skills such as problem-solving, decision-making and rational and clinical thinking. This hybrid strategy can be implemented in many different ways that may fit the faculty goals—e.g., using Flipped Classrooms [126,127]—and this can be done as early as the preclinical years [27].

4.2. What Are the Advantages and Disadvantages of CBL?

Theoretical knowledge is important for any medical professional, but it is also crucial to establish a teaching program that can provide future doctors with other intersocial and clinical skills necessary for the practice of medicine. CBL provides a teaching method that relies on students taking an active role to discuss and resolve clinical cases. This teaching method stimulates students to develop their clinical judgement, which will help them to incorporate this clinical knowledge during their medical careers, resulting in better clinical diagnostic decisions and treatment planning. Adequate clinical decisions can have a relevant socioeconomic impact on patients and the healthcare system by avoiding unnecessary complementary diagnostic testing and inadequate treatments. CBL may result in better economic perception, allowing them to understand how their decisions
will financially impact their patients and the healthcare system, and can thus promote the financial responsibility required for clinical decision-making [39,103]. Although the effect size was not significantly higher in CBL groups for critical thinking, all studies showed a significant and large effect size compared to the control groups, suggesting that critical thinking can be enhanced with CBL.

CBL results in enhanced interest and motivation among medical students, as shown by our meta-analysis. The higher interest and motivation in the dynamic learning process of CBL allows students to be more focused during lectures, thus maximizing their learning potential and resulting in higher class attendance [98,128]. Higher interest and motivation should reflect higher satisfaction with the teaching method. Although the pooled effect was not significant for self-learning and satisfaction outcomes, several subgroups showed a significantly superior effect size for the CBL group, suggesting that satisfaction with the teaching method can be improved with CBL.

Perception of learning efficacy is frequently cited as superior in the CBL groups compared to the traditional teaching method groups [11,90,97,129]. Although this outcome could not be included in our meta-analysis, this may have been reflected in the CBL group achieving significantly higher exam scores.

CBL can also improve bilateral perceptions of this teaching method between students and faculty. Teachers often report that during the discussion of clinical cases, they feel the students are closer, more active and confident in the classes [31,91,130–133]. This may be attributed to the way CBL is implemented, relying on open questions [11,31,94,134], which leaves the students more comfortable and confident to participate in the clinical discussion [27,109,132,134]. Our meta-analysis could not analyze these theoretical advantages due to the scarcity of studies that evaluated student–faculty perceptions and the heterogeneous methods applied by studies that collected these outcomes.

Interpeer collaboration in team work is also an important advantage of CBL. CBL encourages students to share their ideas and knowledge to discuss and resolve clinical cases [41,90,129]. CBL fosters peer interaction, a fundamental skill needed for future medical doctors to work in multidisciplinary teams. Although the pooled effect size of interpeer interaction in team work was not significantly different in our meta-analysis, all but one subgroup showed that this outcome was significantly superior in the CBL groups.

As with any teaching method, CBL is not free of disadvantages. The potential need for more time allotted for classes can be a disadvantage of the CBL teaching method. The time allotted for classes depends on the facilitator profile and how CBL is implemented, but it can be continuously adjusted in response to students’ feedback to provide a more successful implementation and performance [135,136]. Our meta-analysis showed no differences in the time allotted for classes between CBL and didactic lectures. Academic burden outside classes can also be an important disadvantage of CBL. While CBL enhances students’ time management [73,129,137], it can also demand more time from students to prepare for classes [136,138] and result in insufficient time for students to rest between classes [139–141]. These factors can increase stress levels of students and decrease their academic performance [37]. There were not enough studies that evaluated the burden of time consumed before and after classes, and we were not able to perform a meta-analysis on the academic burden. This theoretical disadvantage showed conflicting results across the included studies and the true influence of CBL on out-of-class stress levels remains elusive. While a few studies [92,112] reported no differences in the academic burden or preclass workload, another study [20] even reported less free time consumed in the CBL group. Conversely, another study [37] reported increased stress levels from lack of relaxation and sleeping time in the CBL group, but their outcomes were confounded by the more condensed schedule of the CBL group (reduction by four weeks of the academic semester) compared to the didactic lectures group, which may have resulted in less time to relax and rest during a shorter semester.
4.3. Unanswered Questions and Future Directions

There is a clear need for further studies on this topic with homogeneous research methods that provide extractable and quantifiable data to allow more robust findings to be achieved in future systematic reviews and meta-analyses. Most available comparative studies compare CBL to didactic lectures, but it is important to extend future research to compare other teaching methods that are less represented in our systematic review (TBL, HPS, PBL, tutorial method and independent readings). When there are enough studies comparing these different teaching methods, it will be possible to make a more direct comparison through a network meta-analysis and identify the most effective teaching method(s) for undergraduate medical students.

Future studies should include different stages of the medical curriculum, as the studies included in our systematic review did not cover all curricular areas. It is also important to conduct long-term studies comparing exposure to different teaching methods throughout the undergraduate medical curriculum (from the basic components to the clinics) to understand which methods enhance long-term knowledge retention and improve clinical and interpersonal skills. These long-term studies will help researchers to understand how the maturity of students plays a role in the providing adequate teaching methods and in which curricular year(s) CBL can yield higher efficacy. Our systematic review included studies from different regions of the world, but European and African countries were under-represented. It is important to have data from the other global regions because different countries can represent varying cultural and educational contexts. We encourage future studies on CBL to be conducted within European and African educational contexts.

Although exam scores have an unquestionable impact on students’ academic paths, the perceptions of the teaching method also play an important role. The influence of these perceptions should be evaluated at several timepoints that may be related to CBL—such as in the preparation for classes, during classes and in studying for exams—along with how these perceptions impact the active participation of students and their understanding of the educational content. Our systematic review included only studies that objectively quantified the students’ perceptions, but there is also a significant body of literature that analyzes these perceptions in a qualitative fashion [40,142–147]. Future studies should focus on a mixed-methods approach to combine quantitative and qualitative analysis and reach a deeper understanding of the impact of perception on CBL efficacy and effectiveness.

The assessment of the teachers’ perspective was often overlooked in the included studies. It is crucial to understand how facilitators feel about their less active role in CBL, their tolerance and receptivity to CBL and their vulnerability when discussing clinical cases that may go beyond their specialization, clinical skills and experience. Indeed, some teachers have reported that they felt insecure when transitioning to CBL [47]. These factors can influence the facilitator’s own identity and should also be considered in future studies. There is only limited evidence of the perceptions of teachers in regards to the CBL method. Faculty teaching CBL have reported that a CBL program can be time-intensive [47,49], especially concerning the burden associated with the preparation of clinical cases and classes [148] and coordinating with other facilitator(s) [47]. Notwithstanding, faculty often reported positive feedback [148,149] and welcomed CBL as a complement to conventional methods [150], as it could help them to achieve their objectives [62].

There are some more unanswered questions. Further research should focus on how the clinical cases presented in CBL can impact the efficacy of learning: (i) structured vs. unstructured clinical cases; (ii) use of multimedia (e.g., avatars) or case simulations (with dummies or actors) to complement the clinical vignettes; (iii) adequate duration of classes; and (iv) the benefits of theoretical presentation of the topic (e.g., didactic lectures) before and after the discussion of cases to complement and reinforce the acquired knowledge.
4.4. Limitations

There are some limitations related to the systematic review process. The poor uniformity in the descriptions of CBL across the studies in the literature may have led us to miss some potentially relevant studies during the database searches.

There were also some limitations that were inherent to the original studies. The scarcity of studies that compared CBL with other teaching methodologies (other than didactic lectures) limited the stratification of meta-analysis. The definition of control interventions was also heterogeneous due to the available studies, but we made several sensitivity analyses to account for this confounder. This limitation precluded a head-to-head comparison between CBL and each of the other teaching methodologies. However, a head-to-head comparison between CBL and didactic lectures was possible and is reported in our sensitivity analyses.

Although the CBL is well established and commonly used, was often described using broad pedagogical concepts [11]—i.e., CBL interventions were heterogeneous in terms of conceptual and practical perspectives—which were inherent to each study purpose and specific aim. The poor definition of methods and implementation of CBL and the other teaching methodologies contributed to the difficulty in homogenizing meta-analysis groups, and we thus had to group CBL under a broader umbrella concept of CBL; future studies should allow for a stricter definition of CBL and for the grouping CBL into more homogenous subgroups.

The heterogeneity in how outcomes were collected and the lack of objective quantification of outcome data also precluded us from conducting further meta-analyses on other outcomes that would have been relevant to our systematic review. The variation in how outcomes were collected and the definition and implementation of the teaching methodologies may explain some of the heterogeneity we found, even despite the sensitivity analyses we performed.

The heterogeneity in the included populations, especially related to the curriculum years and areas, may have also led to heterogeneity in the meta-analysis. The imprecision of the reporting of sample sizes (occasionally reported as intervals) decreased the accuracy of the sample size determination; in these cases, where the sample size was reported as a range, we selected the largest sample size.

Studies in medical education are often of a pre–post design which is not ideal for control of all extraneous/confounder factors; this limitation was reflected in a high risk of selection bias.

5. Conclusions

The CBL teaching method significantly enhanced academic performance and improved interest and motivation in undergraduate medical students compared to other teaching methods (especially when compared to didactic lectures or the tutorial method). However, the certainty of evidence was very low and further studies are warranted before a stronger and more definitive conclusion can be drawn. Nonetheless, CBL seems to be superior to or as good as other teaching methods and may be successfully implemented for undergraduate medical students.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci13030238/s1, Table S1—Database search strategy for each database included; Table S2—Risk of bias appraisal criteria defined according to Risk of Bias Assessment tool for Non-randomized Studies (RoBANS) tool; Table S3—Characterization of teaching methodologies according to their models and strategies. Table S4—Scoping summary of outcome measures; Figure S1—Forest plots of included meta-analyses; Figure S1.1—Forest plot comparing exam scores between CBL and other teaching methodologies; Figure S1.2—Forest plot comparing OSCE (objective structured clinical examination) scores between CBL and other teaching methodologies; Figure S1.3—Forest plot comparing self-learning between CBL and other teaching methodologies; Figure S1.4—Forest plot comparing critical thinking between CBL and other teaching methodologies; Figure S1.5—Forest plot comparing satisfaction with teaching method between CBL and other teach-
ing methodologies; Figure S1.6—Forest plot comparing time allotted for classes between CBL and other teaching methodologies; Figure S1.7—Forest plot comparing interest and motivation between CBL and other teaching methodologies; Figure S1.8—Forest plot comparing interpeer interaction in teamwork between CBL and other teaching methodologies; Figure S2—Funnel plots of included meta-analyses; Figure S2.1—Funnel plot of exam score outcomes; Figure S2.2—Trim-and-fill funnel plot of exam score outcomes; Table S5—Sensitivity analyses of each eligible outcome.

Author Contributions: All authors were involved in the conceptualization of the systematic review. D.M. and R.A. were involved in the database searches and data extraction. D.M., R.A. and C.V. performed the data synthesis (meta-analysis), organization and reporting. R.A. and J.A. judged the risk of bias of the studies included in the systematic review. R.A. and C.V. judged the certainty of recommendation using GRADE. J.A., P.C. and J.E.-M. guided and provided advice during all steps of the development of the systematic review, as well as providing critical revision of manuscript content. All authors contributed to drafting and approving the final manuscript prior to submission to the peer-reviewed journal. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare there are no conflict of interest.

References

18. Cendan, J.C.; Silver, M.; Ben-David, K. Changing the student clerkship from traditional lectures to small group case-based sessions benefits the student and the faculty. J. Surg. Educ. 2011, 68, 117–120. [CrossRef]
40. Bowe, C.M.; Voss, J.; Thomas Aretz, H. Case method teaching: An effective approach to integrate the basic and clinical sciences in the preclinical medical curriculum. Med. Teach. 2009, 31, 834–841. [CrossRef]


80. Vandenbroucke, J. Bias in meta-analysis detected by a simple, graphical test. Experts’ views are still needed. *BMJ* 1998, 316, 469.


103. Palappallil, D.S.; Sankar, H.K.; Retnayyan, A.; Radhakrishnan, S. Effectiveness of case-based learning, task-based learning, and didactic lectures on teaching personal drug concept among medical undergraduates. *Natl. J. Physiol. Pharm. Pharmacol.* 2019, 9, 438–444. [CrossRef]


112. Yang, F.; Lin, W.; Wang, Y. Flipped classroom combined with case-based learning is an effective teaching modality in nephrology clerkship. *BMC Med. Educ.* 2021, 21, 276. [CrossRef]


120. Beltrán-Velasco, A.I.; Donoso-González, M.; Clemente-Suárez, V.J. Analysis of perceptual, psychological, and behavioral factors that affect the academic performance of education university students. *Physiol. Behav.* 2021, 238, 113497. [CrossRef]


147. Shetty, J.K.; Begum, S.; Goud, M.B.; Zaki, B. Comparison of didactic lectures and case-based learning in an undergraduate biochemistry course at RAK Medical and Health Sciences University, UAE. *J. Evol. Med. Dent. Sci.* **2016**, *5*, 3212–3216. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.